

Chemokine Alpha-6

Cross Reference to Related Applications

[0001] This application is a divisional of and claims priority under 35 U.S.C. § 120 of U.S. Application No. 09/177,304, filed October 23, 1998, which claims benefit under 35 U.S.C. § 119(e) based on U.S. Provisional Applications Nos. 60/063,387, filed October 24, 1997 and 60/079,245, filed March 25, 1998, all of which are hereby incorporated by reference herein in their entireties.

Field of the Invention

[0002] The present invention relates to a novel human gene encoding a polypeptide which is a member of the chemokine family. More specifically, isolated nucleic acid molecules are provided encoding a human polypeptide named Chemokine Alpha-6, hereinafter referred to as "CK α -6". CK α -6 polypeptides are also provided, as are vectors, host cells and recombinant methods for producing the same. Also provided are diagnostic methods for detecting disorders related to the immune system, and therapeutic methods for treating such disorders. The invention further relates to screening methods for identifying agonists and antagonists of CK α -6 activity.

Background of the Invention

[0003] The ability to control the migration and "trafficking" of various cell types is controlled by a subset of factors, or proteins, among which chemokines are an example.

[0004] Chemokines, also referred to as intercrine cytokines, are a subfamily of structurally and functionally related chemotactic cytokines. These molecules are usually 8-10 kd in size. In general, chemokines exhibit 20% to 75% homology at the amino acid level and are characterized by four conserved cysteine residues that form two disulfide bonds. Based on the arrangement of the first two cysteine residues, chemokines have been classified into two subfamilies, alpha and beta. In the alpha subfamily, the first two cysteines are separated by one amino acid and hence are referred to as the "C-X-C" subfamily. In the beta subfamily, the two cysteines are in an adjacent position and are, therefore, referred to as the "C-C" subfamily. Thus far, at least sixteen different members of this family have been identified in humans.

[0005] The intercrine cytokines exhibit a wide variety of functions. A hallmark feature is their ability to elicit chemotactic migration of distinct cell types, including monocytes, neutrophils, T lymphocytes, basophils and fibroblasts. Many chemokines have proinflammatory activity and are involved in multiple steps during an inflammatory reaction. These activities include stimulation of histamine release, lysosomal enzyme and leukotriene release, increased adherence of target immune cells to endothelial cells, enhanced binding of complement proteins, induced expression of granulocyte adhesion molecules and complement receptors, and respiratory burst. In addition to their involvement in inflammation, certain chemokines have been shown to exhibit other activities. For example, macrophage inflammatory protein 1 (MIP-1) is able to suppress hematopoietic stem cell proliferation, platelet factor-4 (PF-4) is a potent inhibitor of endothelial cell growth, Interleukin-8 (IL-8) promotes proliferation of keratinocytes, and GRO is an autocrine growth factor for melanoma cells.

[0006] In light of the diverse biological activities, it is not surprising that chemokines have been implicated in a number of physiological and disease conditions, including lymphocyte trafficking, wound healing, hematopoietic regulation and immunological disorders such as allergy, asthma and arthritis.

[0008] While members of the C-C branch act predominantly on mononuclear cells and members of the C-X-C branch act predominantly on neutrophils a distinct chemoattractant property cannot be assigned to a chemokine based on this guideline. Some chemokines from one family show characteristics of the other.

[0010] Thus, there is a need for polypeptides that function as regulators of the migration of distinct cell types and of their roles in dysfunction and disease, since disturbances of such regulation may be involved in disorders relating to hemostasis, angiogenesis, tumor metastasis,

cellular migration and ovulation, as well as neurogenesis. Therefore, there is a need for identification and characterization of such human polypeptides which can play a role in detecting, preventing, ameliorating or correcting such disorders.

Summary of the Invention

[0011] The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding at least a portion of the CK α -6 polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 or the complete amino acid sequence encoded by the human cDNA in clone "HFCET92" deposited as plasmid DNA with the American Tissue Culture Collection (ATCC), located at 10801 University Boulevard, Manassas, Virginia 20110-2209 (present address), and assigned ATCC Deposit Number 209300 on September 25, 1997 and ATCC Deposit Number 209643 on February 25, 1998. The nucleotide sequence determined by sequencing the deposited CK α -6 clone (HFCET92), which is shown in Figure 1 (SEQ ID NO:1), contains an open reading frame encoding a complete polypeptide of 84 amino acid residues, including an initiation codon encoding an N-terminal methionine at nucleotide positions 46-48. Nucleic acid molecules of the invention include those encoding the complete amino acid sequence excepting the N-terminal methionine shown in SEQ ID NO:2, or the complete amino acid sequence excepting the N-terminal methionine encoded by a cDNA clone HFCET92, which molecules also can encode additional amino acids fused to the N-terminus of the CK α -6 amino acid sequence.

[0012] The polypeptide of the present invention has amino acid sequence homology to known chemokines, including the conserved C-X-C cysteine pattern characteristic of the alpha subfamily of chemokines beginning with the cysteine at position 22 in SEQ ID NO:2.

[0013] CK α -6 also lacks the ELR motif found in some alpha chemokines immediately preceding the first cysteine residue, which is known to be required for the neutrophil and endothelial cell chemotactic activity as well as the angiogenic activity of IL-8.

[0014] The encoded polypeptide has a predicted leader sequence of about 16, 17, 18, 19, 20 or 21 amino acids (the first 16 of which are underlined in Figure 1); and the amino acid sequence of the predicted mature CK α -6 proteins are also shown in Figure 1 as amino acid residues 17-84, 18-84, 19-84, 20-84, 21-84, and 22-84 (SEQ ID NO:2).

[0015] Thus, one aspect of the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence selected from the group consisting of: (a) a nucleotide

sequence encoding the CK α -6 polypeptide having the complete amino acid sequence in Figure 1 (SEQ ID NO:2); (b) a nucleotide sequence encoding the CK α -6 polypeptide having the complete amino acid sequence in Figure 1 (SEQ ID NO:2) excepting the N-terminal methionine (i.e., positions 2 to 84 of SEQ ID NO:2); (c) a nucleotide sequence encoding the predicted mature CK α -6 polypeptide having the amino acid sequence at positions 16-84, 17-84, 18-84, 19-84, 20-84, 21-84 or 22-84 in Figure 1 (SEQ ID NO:2); (d) a nucleotide sequence encoding a biologically active fragment of the polypeptide of (a); (e) a nucleotide sequence encoding the CK α -6 polypeptide having the complete amino acid sequence encoded by the human cDNA in clone HFCET92; (f) a nucleotide sequence encoding the CK α -6 polypeptide having the complete amino acid sequence excepting the N-terminal methionine encoded by the human cDNA in clone HFCET92; (g) a nucleotide sequence encoding a mature CK α -6 polypeptide encoded by the human cDNA in clone HFCET92; (h) a nucleotide sequence encoding a biologically active fragment of the polypeptide of (e); and (i) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g) or (h), above.

[0016] Further embodiments of the invention include isolated nucleic acid molecules that comprise a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical, to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), (h), or (i) above, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide in (a), (b), (c), (d), (e), (f), (g), (h) or (i), above. This polynucleotide which hybridizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues. An additional nucleic acid embodiment of the invention relates to an isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of a CK α -6 polypeptide having an amino acid sequence in (a), (b), (c), (d), (e), (f), (g) or (h), above.

[0017] The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of CK α -6 polypeptides or peptides by recombinant techniques.

[0018] The invention further provides an isolated CK α -6 polypeptide comprising an amino acid sequence selected from the group consisting of: (a) the amino acid sequence of the full-length CK α -6 polypeptide having the complete amino acid sequence shown in SEQ ID

NO:2 or the complete amino acid sequence encoded by the human cDNA in clone HFCET92; (b) the amino acid sequence of the full-length CK α -6 polypeptide having the complete amino acid sequence shown in SEQ ID NO:2 excepting the N-terminal methionine (i.e., positions 2 to 84 of SEQ ID NO:2) or the complete amino acid sequence excepting the N-terminal methionine encoded by the human cDNA in clone HFCET92; (c) the amino acid sequence of the mature CK α -6 polypeptide having the amino acid sequence shown in SEQ ID NO:2 as residues 17-84, 18-84, 19-84, 20-84, 21-84 or 22-84 or the amino acid sequence of a mature CK α -6 polypeptide encoded by the human cDNA in clone HFCET92; and (d) the amino acid sequence of a biologically active fragment of the polypeptide of (a). The polypeptides of the present invention also include polypeptides having an amino acid sequence at least 80% identical, more preferably at least 90% identical, and still more preferably 95%, 96%, 97%, 98% or 99% identical to those described in (a), (b), (c) or (d) above, as well as polypeptides having an amino acid sequence with at least 90% similarity, and more preferably at least 95% similarity, to those above. Polynucleotides encoding such polypeptides are also provided.

[0019] An additional embodiment of this aspect of the invention relates to a peptide or polypeptide which comprises the amino acid sequence of an epitope-bearing portion of a CK α -6 polypeptide having an amino acid sequence described in (a), (b), (c) or (d), above. Peptides or polypeptides having the amino acid sequence of an epitope-bearing portion of a CK α -6 polypeptide of the invention include portions of such polypeptides with at least six or seven, preferably at least nine, and more preferably at least about 30 amino acids to about 50 amino acids, although epitope-bearing polypeptides of any length up to and including the entire amino acid sequence of a polypeptide of the invention described above also are included in the invention.

[0020] In another embodiment, the invention provides an isolated antibody that binds specifically to a CK α -6 polypeptide having an amino acid sequence described in (a), (b), (c) or (d) above. The invention further provides methods for isolating antibodies that bind specifically to a CK α -6 polypeptide having an amino acid sequence as described herein. Such antibodies are useful diagnostically or therapeutically as described below.

[0021] The invention also provides for pharmaceutical compositions comprising CK α -6 polypeptides, particularly human CK α -6 polypeptides, which may be employed, for instance, to stimulate wound healing, to treat solid tumors, microbial infections, autoimmune diseases, liver

cirrhosis, osteoarthritis, to stimulate neural growth and to treat pulmonary fibrosis. Methods of treating individuals in need of CK α -6 polypeptides are also provided.

[0022] The invention further provides compositions comprising a CK α -6 polynucleotide or a CK α -6 polypeptide for administration to cells *in vitro*, to cells *ex vivo* and to cells *in vivo*, or to a multicellular organism. In certain particularly preferred embodiments of this aspect of the invention, the compositions comprise a CK α -6 polynucleotide for expression of a CK α -6 polypeptide in a host organism for treatment of disease. Particularly preferred in this regard is expression in a human patient for treatment of a dysfunction associated with aberrant endogenous activity of a CK α -6

[0023] The present invention also provides a screening method for identifying compounds capable of enhancing or inhibiting a biological activity of the CK α -6 polypeptide, which involves contacting a receptor polypeptide which is enhanced by the CK α -6 polypeptide with the candidate compound in the presence of a CK α -6 polypeptide, assaying calcium mobilization or chemotactic activity of the cell expressing the receptor in the presence of the candidate compound and of CK α -6 polypeptide, and comparing the receptor activity to a standard level of activity, the standard being assayed when contact is made between the receptor and the CK α -6 polypeptide in the absence of the candidate compound. In this assay, an increase in calcium mobilization or chemotaxis over the standard indicates that the candidate compound is an agonist of CK α -6 activity and a decrease in calcium mobilization or chemotaxis compared to the standard indicates that the compound is an antagonist of CK α -6 activity.

[0024] It has been discovered that CK α -6 is expressed in cDNA libraries derived from adult cerebellum, fetal brain, rejected kidney and bone marrow tissues. Therefore, nucleic acids of the invention are useful as hybridization probes for differential identification of the tissue(s) or cell type(s) present in a biological sample. Similarly, polypeptides and antibodies directed to those polypeptides are useful to provide immunological probes for differential identification of the tissue(s) or cell type(s). In addition, for a number of disorders of the above tissues or cells, particularly of the immune system, significantly higher or lower levels of CK α -6 gene expression may be detected in certain tissues (e.g., cancerous and wounded tissues) or bodily fluids (e.g., serum, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to a "standard" CK α -6 gene expression level, i.e., the CK α -6 expression level in healthy tissue from an individual not having the immune system disorder.

Thus, the invention provides a diagnostic method useful during diagnosis of such a disorder, which involves: (a) assaying CK α -6 gene expression level in cells or body fluid of an individual; (b) comparing the CK α -6 gene expression level with a standard CK α -6 gene expression level, whereby an increase or decrease in the assayed CK α -6 gene expression level compared to the standard expression level is indicative of disorder in the immune system.

[0025] An additional aspect of the invention is related to a method for treating an individual in need of an increased level of CK α -6 activity in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of an isolated CK α -6 polypeptide of the invention or an agonist thereof.

[0026] A still further aspect of the invention is related to a method for treating an individual in need of a decreased level of CK α -6 activity in the body comprising, administering to such an individual a composition comprising a therapeutically effective amount of a CK α -6 antagonist. Preferred antagonists for use in the present invention are CK α -6-specific antibodies.

Brief Description of the Figures

[0027] Figure 1 shows the nucleotide sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) of CK α -6. The predicted leader sequence consists of about 16-21 amino acids (the first 16 of which are underlined).

[0028] Figure 2 shows the regions of identity between the amino acid sequence of the CK α -6 protein and translation product of the human mRNA for chemokine alpha-2 (SEQ ID NO:3), chemokine alpha-3 (SEQ ID NO:4), and chemokine alpha-4 (SEQ ID NO:5), as determined by Megalign (contained in the DNA Star suite) using the clustal method.

[0029] Figure 3 shows an analysis of the CK α -6 amino acid sequence. Alpha, beta, turn and coil regions; hydrophilicity and hydrophobicity; amphipathic regions; flexible regions; antigenic index and surface probability are shown. In the "Antigenic Index - Jameson-Wolf" graph, the positive peaks indicate locations of the highly antigenic regions of the CK α -6 protein, i.e., regions from which epitope-bearing peptides of the invention can be obtained.

Detailed Description

[0030] The present invention provides isolated nucleic acid molecules comprising a polynucleotide encoding a CK α -6 polypeptide having the amino acid sequence shown in SEQ

ID NO:2, which was determined by sequencing cloned cDNAs. The nucleotide sequence shown in Figure 1 (SEQ ID NO:1) was obtained by sequencing cDNA clone: HFCET92. HFCET92 was deposited twice with the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VA 20110-2209. The first deposit was made on September 25, 1997 and was given ATCC Deposit No. 209300. This deposit is comprised of a mixture of 50 cDNA species of which clone HFCET92 is one. The second deposit was made on February 25, 1998 and was given ATCC Deposit No. 209643. This second deposit is a homogeneous population of HFCET92 clones. The deposited HFCET92 clones are contained in the pBluescript SK(-) plasmid (Stratagene, La Jolla, CA).

[0031] The polypeptide of the present invention has amino acid sequence homology to known chemokines, including the conserved C-X-C cysteine pattern characteristic of the alpha subfamily of chemokines beginning with the cysteine at position 22 counting from the amino terminus in SEQ ID NO:2. The CK α -6 protein of the present invention also shares sequence homology with other chemokines including in particular the translation product of the human mRNAs for chemokine alpha-2 (SEQ ID NO:3) described in Patent Cooperation Treaty Application Serial No. US97/04329, chemokine alpha-3 (SEQ ID NO:4) described in Patent Cooperation Treaty Application Serial No. US96/03686, and chemokine alpha-4 (SEQ ID NO:5) described in Patent Cooperation Treaty Application Serial No. US96/14630, as shown in Figure 2.

[0032] Of the known members of the alpha chemokine family, the majority contain an ELR motif (e.g., IL-8, ENA-78, GCP2, GRO- α , PBP, CTAP-III and NAP-2) and others lack the ELR motif (IP-10, PF4 and MIG). CK α -6 lacks the ELR motif immediately preceding the first cysteine residue. It has been clearly shown that this ELR motif is required for the neutrophil and endothelial cell chemotactic activity as well as the angiogenic activity of IL-8 (Strieter *et al.*, *J. Biol. Chem.*, 270:27348-27357 (1995)). In addition, it has been shown that the ELR-C-X-C chemokines have *in vitro* and *in vivo* angiogenic activity, whereas the C-X-C chemokines lacking the ELR motif are actually angiostatic (Strieter *et al.*, *supra*; Angiolillo *et al.*, *J. Exp. Med.*, 182:155-162 (1995); and Koch *et al.*, *Science*, 258:1798-1801 (1992)). In terms of a possible role of such factors in tumor angiogenesis, Smith *et al.*, *J. Exp. Med.*, 179:1409-1415 (1994), has reported increased IL-8 levels in bronchogenic carcinoma tumor tissues which appear to be produced from the tumor cells.

Nucleic Acid Molecules

[0033] Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequencer (such as the Model 373 from Applied Biosystems, Inc., Foster City, CA), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were predicted by translation of a DNA sequence determined as above. Therefore, as is known in the art for any DNA sequence determined by this automated approach, any nucleotide sequence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide sequence of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a determined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded by the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

[0034] By "nucleotide sequence" of a nucleic acid molecule or polynucleotide is intended, for a DNA molecule or polynucleotide, a sequence of deoxyribonucleotides, and for an RNA molecule or polynucleotide, the corresponding sequence of ribonucleotides (A, G, C and U), where each thymidine deoxyribonucleotide (T) in the specified deoxyribonucleotide sequence is replaced by the ribonucleotide uridine (U).

[0035] Using the information provided herein, such as the nucleotide sequence in Figure 1 (SEQ ID NO:1), a nucleic acid molecule of the present invention encoding a CK α -6 polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Illustrative of the invention, the nucleic acid molecule described in Figure 1 (SEQ ID NO:1) was discovered in cDNA libraries derived from brain and immune-system tissues. Clone HFCET92, for example, was isolated from a fetal brain cDNA library.

[0036] Additional clones of the same gene were also identified in cDNA libraries from the following human tissues: adult brain, rejected spleen and bone marrow.

[0037] The determined nucleotide sequence of the CK α -6 cDNA of Figure 1 (SEQ ID NO:1) contains an open reading frame encoding a protein of 84 amino acid residues, with an

initiation codon at nucleotide positions 46-48 of the nucleotide sequence in Figure 1 (SEQ ID NO:1).

[0038] As one of ordinary skill would appreciate, due to the possibility of sequencing errors discussed above, the actual complete CK α -6 polypeptide encoded by the deposited cDNA, which comprises about 84 amino acids, may be somewhat longer or shorter. More generally, the actual open reading frame may be anywhere in the range of ± 20 amino acids, more likely in the range of ± 10 amino acids, of that predicted from the methionine codon at the N-terminus shown in Figure 1 (SEQ ID NO:1).

Leader and Mature Sequences

[0039] The amino acid sequence of the complete CK α -6 protein includes a leader sequence and a mature protein, as shown in SEQ ID NO:2. More in particular, the present invention provides nucleic acid molecules encoding a mature form of the CK α -6 protein. Thus, according to the signal hypothesis, once export of the growing protein chain across the rough endoplasmic reticulum has been initiated, proteins secreted by mammalian cells have a signal or secretory leader sequence which is cleaved from the complete polypeptide to produce a secreted "mature" form of the protein. Most mammalian cells and even insect cells cleave secreted proteins with the same specificity. However, in some cases, cleavage of a secreted protein is not entirely uniform, which results in two or more mature species of the protein. Further, it has long been known that the cleavage specificity of a secreted protein is ultimately determined by the primary structure of the complete protein, that is, it is inherent in the amino acid sequence of the polypeptide. Therefore, the present invention provides a nucleotide sequence encoding the mature CK α -6 polypeptide having the amino acid sequence encoded by the cDNA clone identified as HFCET92. By the "mature CK α -6 polypeptide having the amino acid sequence encoded by the cDNA clone HFCET92" is meant the mature form(s) of the CK α -6 protein produced by expression in a mammalian cell (e.g., COS cells, as described below) of the complete open reading frame encoded by the human DNA sequence of the clone contained in the deposited vector.

[0040] In addition, methods for predicting whether a protein has a secretory leader as well as the cleavage point for that leader sequence are available. For instance, the method of McGeoch (*Virus Res.* 3:271-286 (1985)) uses the information from a short N-terminal charged region and a subsequent uncharged region of the complete (uncleaved) protein. The method of von Heinje (*Nucleic Acids Res.* 14:4683-4690 (1986)) uses the information from the residues

surrounding the cleavage site, typically residues -13 to +2 where +1 indicates the amino terminus of the mature protein. The accuracy of predicting the cleavage points of known mammalian secretory proteins for each of these methods is in the range of 75-80% (von Heinje, *supra*). However, the two methods do not always produce the same predicted cleavage point(s) for a given protein.

[0041] In the present case, the deduced amino acid sequence of the complete CK α -6 polypeptide was analyzed by a computer program "PSORT", available from Dr. Kenta Nakai of the Institute for Chemical Research, Kyoto University (see K. Nakai and M. Kanehisa, *Genomics* 14:897-911 (1992)), and "SignalP", described in Nielsen et al., *Protein Engineering* 10:1-6 (1997), which are both expert systems for predicting the cellular location of a protein based on the amino acid sequence. As part of these computational prediction of localization, the methods of McGeoch and von Heinje are incorporated." The computation analysis above predicted one potential cleavage site within the complete amino acid sequence shown in SEQ ID NO:2; that is, between residues 21 and 22 in Figure 1 (SEQ ID NO:2). Thus, an especially preferred mature polypeptide comprises residues 22 to 84 in Figure 1 (SEQ ID NO:2). Because computer algorithms may not always accurately predict the correct cleavage site used in nature and because cleavage sites are known to vary from one organism to the next, the inventors have designated several mature polypeptides including: residues 17 to 84, residues 18 to 84, residues 19 to 84, residues 20 to 84, residues 21 to 84 and residues 22 to 82 as those residue position numbers are shown in Figure 1 (SEQ ID NO:2).

[0042] As indicated, nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced synthetically. The DNA may be double-stranded or single-stranded. Single-stranded DNA or RNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

[0043] By "isolated" nucleic acid molecule(s) is intended a nucleic acid molecule, DNA or RNA, which has been removed from its native environment. For example, recombinant DNA molecules contained in a vector are considered isolated for the purposes of the present invention. Further examples of isolated DNA molecules include recombinant DNA molecules maintained in heterologous host cells or purified (partially or substantially) DNA molecules in solution. Isolated RNA molecules include *in vivo* or *in vitro* RNA transcripts of the DNA molecules of the present invention. Isolated nucleic acid molecules according to the present invention further include such molecules produced synthetically.

[0051] In addition, the following nucleic acid molecules having nucleotide sequences related to extensive portions of SEQ ID NO:1 have been identified: sequence "HCEOU59R" (SEQ ID NO:14) from cDNA clone "HCEOU59" and sequence "HFCET92RB" (SEQ ID NO:19) from cDNA clone "HFCET92RB."

[0052] The following public expressed sequence tags ("ESTs"), which relate to portions of SEQ ID NO:1, have also been identified: GenBank Accession No. AA410918 (SEQ ID NO:6); GenBank Accession No. AA419299 (SEQ ID NO:7); GenBank Accession No. AA411042 (SEQ ID NO:8); GenBank Accession No. AA410950 (SEQ ID NO:9); GenBank Accession No. AA325795 (SEQ ID NO:10); GenBank Accession No. D20974 (SEQ ID NO:11); and GenBank Accession No. AA700891 (SEQ ID NO:12). Each of the foregoing GenBank references are incorporated herein by reference in their entireties. Nucleic acid molecules of the invention preferably do not comprise a sequence of any individual one (or more) of the foregoing EST sequences. The invention may in addition preferably not contain within its scope any one (or more) of the complete nucleic acid sequences of the clones (including vector sequences) from which the EST sequences were obtained. Nucleic acid molecules of the invention preferably do not comprise the complete sequence of any of the pBS vectors (Stratagene) and/or the pT7T3D vector (Pharmacia).

[0053] Further, the invention includes a polynucleotide comprising any portion of at least about 17 nucleotides, preferably at least about 20, 25, or 30, and most preferably at least about 50 nucleotides, of SEQ ID NO:1 from nucleotide 46 to 297, provided that said polynucleotide is not an EST shown in SEQ ID NOS:6-12 and 14.

[0054] More generally, by a fragment of an isolated nucleic acid molecule having the nucleotide sequence of the deposited cDNA or the nucleotide sequence shown in Figure 1 (SEQ ID NO:1) is intended fragments at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt in length which are useful as diagnostic probes and primers as discussed herein. Of course, larger fragments 50-1,000 nt in length are also useful according to the present invention as are fragments corresponding to most, if not all, of the nucleotide sequence of the deposited cDNA or as shown in Figure 1 (SEQ ID NO:1). By a fragment at least 20 nt in length, for example, is intended fragments which include 20 or more contiguous bases from the nucleotide sequence of the deposited cDNA or the nucleotide sequence as shown in Figure 1 (SEQ ID NO:1). Preferred nucleic acid fragments of the present invention include nucleic acid molecules

encoding epitope-bearing portions of the CK α -6 polypeptide as identified from the Jameson-Wolf antigenic index shown in Figure 3 and described in more detail below.

[0055] In another aspect, the invention provides an isolated nucleic acid molecule comprising a polynucleotide which hybridizes under stringent hybridization conditions to a portion of the polynucleotide in a nucleic acid molecule of the invention described above, for instance, the cDNA clone HFCET92, preferably the coding portion. By "stringent hybridization conditions" is intended overnight incubation at 42° C in a solution comprising: 50% formamide, 5x SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 μ g/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1x SSC at about 65° C.

[0056] By a polynucleotide which hybridizes to a "portion" of a polynucleotide is intended a polynucleotide (either DNA or RNA) hybridizing to at least about 15 nucleotides (nt), and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably about 30-70 (e.g., 50) nt of the reference polynucleotide. These are useful as diagnostic probes and primers as discussed above and in more detail below.

[0057] By a portion of a polynucleotide of "at least 20 nt in length," for example, is intended 20 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide (e.g., the deposited cDNA or the nucleotide sequence as shown in Figure 1 (SEQ ID NO:1)). Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3' terminal poly(A) tract of the CK α -6 cDNA shown in Figure 1 (SEQ ID NO:1)), or to a complementary stretch of T (or U) residues, would not be included in a polynucleotide of the invention used to hybridize to a portion of a nucleic acid of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone).

[0058] As indicated, nucleic acid molecules of the present invention which encode a CK α -6 polypeptide may include, but are not limited to those encoding the amino acid sequence of the mature polypeptide, by itself; and the coding sequence for the mature polypeptide and additional sequences, such as those encoding the about 21 amino acid leader or secretory sequence, such as a pre-, or pro- or prepro- protein sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences.

[0059] Also encoded by nucleic acids of the invention are the above protein sequences together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences

that play a role in transcription, mRNA processing, including splicing and polyadenylation signals, for example - ribosome binding and stability of mRNA; an additional coding sequence which codes for additional amino acids, such as those which provide additional functionalities.

[0060] Thus, the sequence encoding the polypeptide may be fused to a marker sequence, such as a sequence encoding a peptide which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, CA, 91311), among others, many of which are commercially available. As described in Gentz *et al.*, *Proc. Natl. Acad. Sci. USA* 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. The "HA" tag is another peptide useful for purification which corresponds to an epitope derived from the influenza hemagglutinin protein, which has been described by Wilson *et al.*, *Cell* 37: 767 (1984). As discussed below, other such fusion proteins include the CK α -6 fused to Fc at the N- or C-terminus.

Variant and Mutant Polynucleotides

[0061] The present invention further relates to variants of the nucleic acid molecules of the present invention, which encode portions, analogs or derivatives of the CK α -6 protein. Variants may occur naturally, such as a natural allelic variant. By an "allelic variant" is intended one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. *Genes II*, Lewin, B., ed., John Wiley & Sons, New York (1985). Non-naturally occurring variants may be produced using art-known mutagenesis techniques.

[0062] Such variants include those produced by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding regions, non-coding regions, or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the CK α -6 protein or portions thereof. Also especially preferred in this regard are conservative substitutions.

[0063] Most highly preferred are nucleic acid molecules encoding the mature protein having the amino acid sequence shown in SEQ ID NO:2 or the mature CK α -6 amino acid sequence encoded by the deposited cDNA clone either of which may be modified so as to

encode an amino terminal methionine or other N- or C-terminal fusion peptides or polypeptides.

[0064] Further embodiments include an isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical to a polynucleotide selected from the group consisting of: (a) a nucleotide sequence encoding the CK α -6 polypeptide having the complete amino acid sequence in SEQ ID NO:2; (b) a nucleotide sequence encoding the CK α -6 polypeptide having the complete amino acid sequence in SEQ ID NO:2 excepting the N-terminal methionine (i.e., positions 2 to 84 of SEQ ID NO:2); (c) a nucleotide sequence encoding the predicted mature CK α -6 polypeptide having the amino acid sequence at positions 17-84, 18-84, 19-84, 20-84, 21-84, and/or 22-84 in SEQ ID NO:2; (d) a nucleotide sequence encoding a biologically active fragment of the polypeptide of (a); (e) a nucleotide sequence encoding the CK α -6 polypeptide having the complete amino acid sequence encoded by the human cDNA in clone HFCET92; (f) a nucleotide sequence encoding the CK α -6 polypeptide having the complete amino acid sequence excepting the N-terminal methionine encoded by the human cDNA in clone HFCET92; (g) a nucleotide sequence encoding a mature CK α -6 polypeptide encoded by the human cDNA in clone HFCET92; (h) a nucleotide sequence encoding a biologically active fragment of the polypeptide of (e); and (i) a nucleotide sequence complementary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g) or (h) above.

[0065] Further embodiments of the invention include isolated nucleic acid molecules that comprise a polynucleotide having a nucleotide sequence at least 90% identical, and more preferably at least 95%, 96%, 97%, 98% or 99% identical, to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), (h) or (i) above, or a polynucleotide which hybridizes under stringent hybridization conditions to a polynucleotide in (a), (b), (c), (d), (e), (f), (g), (h) or (i), above. This polynucleotide which hybridizes does not hybridize under stringent hybridization conditions to a polynucleotide having a nucleotide sequence consisting of only A residues or of only T residues. An additional nucleic acid embodiment of the invention relates to an isolated nucleic acid molecule comprising a polynucleotide which encodes the amino acid sequence of an epitope-bearing portion of a CK α -6 polypeptide having an amino acid sequence in (a), (b), (c), (d), (e), (f), (g) or (h) above.

[0066] The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the

recombinant vectors, as well as to methods of making such vectors and host cells and for using them for production of CK α -6 polypeptides or peptides by recombinant techniques.

[0067] By a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence encoding a CK α -6 polypeptide is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the CK α -6 polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. These mutations of the reference sequence may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence.

[0068] As a practical matter, whether any particular nucleic acid molecule is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the nucleotide sequence shown in Figure 1 or to the nucleotides sequence of the deposited cDNA clone can be determined conventionally using known computer programs such as the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). Bestfit uses the local homology algorithm of Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981), to find the best segment of homology between two sequences. When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference nucleotide sequence and that gaps in homology of up to 5% of the total number of nucleotides in the reference sequence are allowed.

[0069] In a preferred method for determining % identity between two polynucleotide sequences, by a polynucleotide having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the polynucleotide is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the

reference nucleotide sequence encoding the polypeptide. In other words, to obtain a polynucleotide having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence shown in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

[0070] As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB alignment of DNA sequences to calculate percent identity are: Matrix=Unitary, k-tuple=4, Mismatch Penalty=1, Joining Penalty=30, Randomization Group Length=0, Cutoff Score=1, Gap Penalty=5, Gap Size Penalty 0.05, Window Size=500 or the length of the subject nucleotide sequence, whichever is shorter.

[0071] If the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This corrected score is what is used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the FASTDB alignment, which are not matched/aligned with

the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

[0072] For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the FASTDB alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the FASTDB program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

[0073] The present application is directed to nucleic acid molecules at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1) or to the nucleic acid sequence of the deposited cDNA, irrespective of whether they encode a polypeptide having CK α -6 activity. This is because even where a particular nucleic acid molecule does not encode a polypeptide having CK α -6 activity, one of skill in the art would still know how to use the nucleic acid molecule, for instance, as a hybridization probe or a polymerase chain reaction (PCR) primer. Uses of the nucleic acid molecules of the present invention that do not encode a polypeptide having CK α -6 activity include, inter alia, (1) isolating the CK α -6 gene or allelic variants thereof in a cDNA library; (2) *in situ* hybridization (e.g., "FISH") to metaphase chromosomal spreads to provide precise chromosomal location of the CK α -6 gene, as described in Verma et al., Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York (1988); and Northern Blot analysis for detecting CK α -6 mRNA expression in specific tissues.

[0074] Preferred, however, are nucleic acid molecules having sequences at least 90%, 95%, 96%, 97%, 98% or 99% identical to the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1) or to the nucleic acid sequence of the deposited cDNA which do, in fact, encode a polypeptide having CK α -6 protein activity. By "a polypeptide having CK α -6 activity" is

intended polypeptides exhibiting activity similar, but not necessarily identical, to an activity of other alpha chemokines such as calcium mobilization in immune system tissues and/or chemotaxis of immune-system tissues. Such assays are well known in the art. See, for example, see Peters *et al.*, *Immun. Today* 17:273 (1996); Young *et al.*, *J. Exp. Med.* 182:1111 (1995); Caux *et al.*, *Nature* 390:258 (1992); and Santiago-Schwarz *et al.*, *Adv. Exp. Med. Biol.* 378:7 (1995).

[0075] CK α -6 protein modulates calcium mobilization in, and chemotaxis of, immune system cells in a dose-dependent manner in the above-described assays. Thus, "a polypeptide having CK α -6 protein activity" includes polypeptides that also exhibit any of the same activities in the above-described assays in a dose-dependent manner. Although the degree of dose-dependent activity need not be identical to that of the CK α -6 protein, preferably, "a polypeptide having CK α -6 protein activity" will exhibit substantially similar dose-dependence in a given activity as compared to the CK α -6 protein (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity relative to the reference CK α -6 protein).

[0076] Like other CXC chemokines, CK α -6 exhibits activity on leukocytes including for example monocytes, lymphocytes and neutrophils. For this reason CK α -6 is active in directing the proliferation, differentiation and migration of these cell types. Such activity is useful for immune enhancement or suppression, myeloprotection, stem cell mobilization, acute and chronic inflammatory control and treatment of leukemia. Assays for measuring such activity are known in the art. For example, see Peters *et al.*, *Immun. Today* 17:273 (1996); Young *et al.*, *J. Exp. Med.* 182:1111 (1995); Caux *et al.*, *Nature* 390:258 (1992); and Santiago-Schwarz *et al.*, *Adv. Exp. Med. Biol.* 378:7 (1995), all of which are incorporated by reference herein.

[0077] Of course, due to the degeneracy of the genetic code, one of ordinary skill in the art will immediately recognize that a large number of the nucleic acid molecules having a sequence at least 90%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequence of the deposited cDNA or the nucleic acid sequence shown in Figure 1 (SEQ ID NO:1) will encode a polypeptide "having CK α -6 protein activity." In fact, since degenerate variants of these nucleotide sequences all encode the same polypeptide, this will be clear to the skilled artisan even without performing the above described comparison assay. It will be further recognized in the art that, for such nucleic acid molecules that are not degenerate variants, a reasonable number will also encode a polypeptide having CK α -6 protein activity. This is

because the skilled artisan is fully aware of amino acid substitutions that are either less likely or not likely to significantly effect protein function (e.g., replacing one aliphatic amino acid with a second aliphatic amino acid), as further described below.

Vectors and Host Cells

[0078] The present invention also relates to vectors which include the isolated DNA molecules of the present invention, host cells which are genetically engineered with the recombinant vectors, and the production of CK α -6 polypeptides or fragments thereof by recombinant techniques. The vector may be, for example, a phage, plasmid, viral or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

[0079] The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

[0080] The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac*, *trp*, *phoA* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

[0081] As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, 293 and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

[0082] Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc., *supra*; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and p

[0083] Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis *et al.*, *Basic Methods In Molecular Biology* (1986).

[0084] The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to stabilize and purify proteins. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is thoroughly advantageous for use in therapy and diagnosis and thus results, for example, in improved pharmacokinetic properties (EP-A 0232 262). On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified in the advantageous manner described. This is the case when Fc portion proves to be a hindrance to use in therapy and diagnosis, for example when the fusion protein is to be used as antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify

antagonists of hIL-5. See, D. Bennett *et al.*, *J. Molecular Recognition* 8:52-58 (1995) and K. Johanson *et al.*, *J. Biol. Chem.* 270:9459-9471 (1995).

[0085] The CK α -6 protein can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Polypeptides of the present invention include: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

Polypeptides and Fragments

[0086] The invention further provides an isolated CK α -6 polypeptide having the amino acid sequence encoded by the deposited human cDNA in clone HFCET92, or the amino acid sequence in SEQ ID NO:2, or a peptide or polypeptide comprising a portion of the above polypeptides.

Variant and Mutant Polypeptides

[0087] To improve or alter the characteristics of CK α -6 polypeptides, protein engineering may be employed. Recombinant DNA technology known to those skilled in the art can be used to create novel mutant proteins or "muted" proteins including single or multiple amino acid substitutions, deletions, additions or fusion proteins. Such modified polypeptides can show,

e.g., enhanced activity or increased stability. In addition, they may be purified in higher yields and show better solubility than the corresponding natural polypeptide, at least under certain purification and storage conditions.

N-Terminal and C-Terminal Deletion Mutants

[0088] For instance, for many proteins, including the extracellular domain of a membrane associated protein or the mature form(s) of a secreted protein, it is known in the art that one or more amino acids may be deleted from the N-terminus or C-terminus without substantial loss of biological function. For instance, Ron et al., *J. Biol. Chem.*, 268:2984-2988 (1993) reported modified KGF proteins that had heparin binding activity even if 3, 8, or 27 amino-terminal amino acid residues were missing. In the present case, since the protein of the invention is a member of the chemokine polypeptide family, deletions of N-terminal amino acids up to the first "Cys" required for formation of a disulfide bridge may retain some biological activity such as receptor binding or modulation of target cell activities. Polypeptides having further N-terminal deletions including the cysteine residue at position 22 in Figure 1 (SEQ ID NO:2) would not be expected to retain such biological activities because it is known that this residue in a chemokine-related polypeptide is required for forming a disulfide bridge to provide structural stability which is needed for receptor binding and signal transduction.

[0089] However, even if deletion of one or more amino acids from the N-terminus of a protein results in modification or loss of one or more biological functions of the protein, other biological activities may still be retained. Thus, the ability of the shortened protein to induce and/or bind to antibodies which recognize the complete or mature form of the protein generally will be retained when less than the majority of the residues of the complete or mature protein are removed from the N-terminus. Whether a particular polypeptide lacking N-terminal residues of a complete protein retains such immunologic activities can readily be determined by routine methods described herein and otherwise known in the art.

[0090] Accordingly, the present invention further provides polypeptides having one or more residues deleted from the amino terminus of the amino acid sequence of the CK α -6 shown in SEQ ID NO:2, up to the cysteine residue at position number 22, and polynucleotides encoding such polypeptides. In particular, the present invention provides polypeptides comprising the amino acid sequence of residues n-84 of SEQ ID NO:2, where n is an integer in the range of 17 to 22, and Cys-22 is the position of the first residue from the N-terminus of the

first residue from the C- terminus of the complete CK α -6 polypeptide (shown in SEQ ID NO:2) believed to be required for receptor binding of the CK α -6 protein.

[0095] More in particular, the invention provides polynucleotides encoding polypeptides having the amino acid sequence of residues of 17 to 68, 17 to 69, 17 to 70, 17 to 71, 17 to 72, 17 to 73, 17 to 74, 17 to 75, 17 to 76, 17 to 77, 17 to 78, 17 to 79, 17 to 80, 17 to 81, 17 to 82, 17 to 83 and 17 to 84, all of SEQ ID NO:2. Polynucleotides encoding these polypeptides also are provided.

[0096] The invention also provides polypeptides having one or more amino acids deleted from both the amino and the carboxyl termini of the full-length polypeptide, which may be described generally as having residues n-m of SEQ ID NO:2, where n and m are integers as described above.

[0097] Also included are a nucleotide sequence encoding a polypeptide consisting of a portion of the complete CK α -6 amino acid sequence encoded by the cDNA clone contained in ATCC Deposit No. 209231, where this portion excludes from 16 to about 21 amino acids from the amino terminus of the full-length CK α -6 polypeptide encoded by the cDNA clone HFCET92, or from 1 to about 16 amino acids from the carboxy terminus, or any combination of the above amino terminal and carboxy terminal deletions, of the full-length protein encoded by HFCET92. Polynucleotides encoding all of the above deletion mutant polypeptide forms also are provided.

Other Mutants

[0098] In addition to terminal deletion forms of the protein discussed above, it also will be recognized by one of ordinary skill in the art that some amino acid sequences of the CK α -6 polypeptide can be varied without significant effect of the structure or function of the protein. If such differences in sequence are contemplated, it should be remembered that there will be critical areas on the protein which determine activity.

[0099] Thus, the invention further includes variations of the CK α -6 polypeptide which show substantial CK α -6 polypeptide activity or which include regions of CK α -6 protein such as the protein portions discussed below. Such mutants include deletions, insertions, inversions, repeats, and type substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie, J. U. et al., "Deciphering the Message in Protein Sequences: Tolerance to Amino Acid Substitutions," *Science* 247:1306-1310 (1990), wherein

TABLE 1. Conservative Amino Acid Substitutions.

Aromatic	Phenylalanine
	Tryptophan
	Tyrosine
Hydrophobic	Leucine
	Isoleucine
	Valine
Polar	Glutamine
	Asparagine
Basic	Arginine
	Lysine
	Histidine
Acidic	Aspartic Acid
	Glutamic Acid
Small	Alanine
	Serine
	Threonine
	Methionine
	Glycine

[0103] Amino acids in the CK α -6 protein of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, *Science* 244:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity such as receptor binding or *in vitro* or *in vitro* proliferative activity.

[0104] Of special interest are substitutions of charged amino acids with other charged or neutral amino acids which may produce proteins with highly desirable improved

characteristics, such as less aggregation. Aggregation may not only reduce activity but also be problematic when preparing pharmaceutical formulations, because aggregates can be immunogenic (Pinckard *et al.*, *Clin. Exp. Immunol.* 2:331-340 (1967); Robbins *et al.*, *Diabetes* 36: 838-845 (1987); Cleland *et al.*, *Crit. Rev. Therapeutic Drug Carrier Systems* 10:307-377 (1993).

[0105] Replacement of amino acids can also change the selectivity of the binding of a ligand to cell surface receptors. For example, Ostade *et al.*, *Nature* 361:266-268 (1993) describes certain mutations resulting in selective binding of TNF- α to only one of the two known types of TNF receptors. Sites that are critical for ligand-receptor binding can also be determined by structural analysis such as crystallization, nuclear magnetic resonance or photoaffinity labeling (Smith *et al.*, *J. Mol. Biol.* 224:899-904 (1992) and de Vos *et al. Science* 255:306-312 (1992)).

[0106] As described above CK α -6 lacks an ELR motif. By making a specific mutation in CK α -6 to include an ELR motif in the position where such a motif is typically found in related chemokines, CK α -6 will act as an antagonist, thus possessing angiogenic activity. Accordingly, polypeptides of the present invention include CK α -6-ELR mutants. Such CK α -6-ELR mutants are comprised of the full-length or preferably the mature CK α -6 protein. Further, since CK α -6 is a member of the chemokine-related protein family, to modulate rather than completely eliminate biological activities of CK α -6 preferably mutations are made in residues not conserved in all members of the chemokine family. As is known in the art, the four spatially conserved cysteines present in all chemokines, positions 22, 24, 48 and 68 in CK α -6, are required for the formation of two disulfide bridges. Thus, it is preferable not to alter any of the four cysteine residues located at positions 22, 24, 48 and 68 in SEQ ID NO:2. Also forming part of the present invention are isolated polynucleotides comprising nucleic acid sequences which encode the above CK α -6 mutants.

[0107] The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of the CK α -6 polypeptide can be substantially purified by the one-step method described in Smith and Johnson, *Gene* 67:31-40 (1988). Polypeptides of the invention also can be purified from natural or recombinant sources using anti-CK α -6 antibodies of the invention in methods which are well known in the art of protein purification.

1. The first step is to identify the problem. This involves understanding the current situation and what needs to be changed.

1. The first step is to identify the problem. This involves understanding the current situation and what needs to be changed.

1. The first step is to identify the problem. This involves understanding the current situation and what needs to be changed.

1. The first step is to identify the problem. This involves understanding the current situation and what needs to be changed.

having an amino acid sequence at least 95% identical to a reference amino acid sequence, up to 5% of the amino acid residues in the reference sequence may be deleted or substituted with another amino acid, or a number of amino acids up to 5% of the total amino acid residues in the reference sequence may be inserted into the reference sequence. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

[0112] As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequence shown in SEQ ID NO:2 or to the amino acid sequence encoded by deposited cDNA clone can be determined conventionally using known computer programs such the Bestfit program (Wisconsin Sequence Analysis Package, Version 8 for Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711). When using Bestfit or any other sequence alignment program to determine whether a particular sequence is, for instance, 95% identical to a reference sequence according to the present invention, the parameters are set, of course, such that the percentage of identity is calculated over the full length of the reference amino acid sequence and that gaps in homology of up to 5% of the total number of amino acid residues in the reference sequence are allowed.

[0113] The following is a preferred method for determining "% identity" as between two polypeptide sequences: By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, (indels) or substituted with another amino acid. These alterations of the reference sequence may occur at the amino or carboxy terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

[0114] As a practical matter, whether any particular polypeptide is at least 90%, 95%, 96%, 97%, 98% or 99% identical to, for instance, the amino acid sequences shown in Table 1

or to the amino acid sequence encoded by deposited DNA clone can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the FASTDB computer program based on the algorithm of Brutlag et al. (Comp. App. Biosci. (1990) 6:237-245). In a sequence alignment the query and subject sequences are either both nucleotide sequences or both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a FASTDB amino acid alignment are: Matrix=PAM 0, k-tuple=2, Mismatch Penalty=1, Joining Penalty=20, Randomization Group Length=0, Cutoff Score=1, Window Size=sequence length, Gap Penalty=5, Gap Size Penalty=0.05, Window Size=500 or the length of the subject amino acid sequence, whichever is shorter.

[0115] If the subject sequence is shorter than the query sequence due to N- or C-terminal deletions, not because of internal deletions, a manual correction must be made to the results. This is because the FASTDB program does not account for N- and C-terminal truncations of the subject sequence when calculating global percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the FASTDB sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above FASTDB program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what is used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

[0116] For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the FASTDB alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C- termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score

calculated by the FASTDB program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by FASTDB is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the FASTDB alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are to be made for the purposes of the present invention.

[0117] The polypeptide of the present invention could be used as a molecular weight marker on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art.

[0118] As described in detail below, the polypeptides of the present invention can also be used to raise polyclonal and monoclonal antibodies, which are useful in assays for detecting CK α -6 protein expression as described below or as agonists and antagonists capable of enhancing or inhibiting CK α -6 protein function. Further, such polypeptides can be used in the yeast two-hybrid system to "capture" CK α -6 protein binding proteins which are also candidate agonists and antagonists according to the present invention. The yeast two hybrid system is described in Fields and Song, *Nature* 340:245-246 (1989).

Epitope-Bearing Portions

[0119] In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention. The epitope of this polypeptide portion is an immunogenic or antigenic epitope of a polypeptide of the invention. An "immunogenic epitope" is defined as a part of a protein that elicits an antibody response when the whole protein is the immunogen. On the other hand, a region of a protein molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a protein generally is less than the number of antigenic epitopes. See, for instance, Geysen *et al.*, *Proc. Natl. Acad. Sci. USA* 81:3998- 4002 (1983).

[0120] As to the selection of peptides or polypeptides bearing an antigenic epitope (i.e., that contain a region of a protein molecule to which an antibody can bind), it is well known in that art that relatively short synthetic peptides that mimic part of a protein sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked protein. See,

for instance, Sutcliffe, J. G., Shinnick, T. M., Green, N. and Learner, R. A. (1983) "Antibodies that react with predetermined sites on proteins," *Science*, 219:660-666. Peptides capable of eliciting protein-reactive sera are frequently represented in the primary sequence of a protein, can be characterized by a set of simple chemical rules, and are confined neither to immunodominant regions of intact proteins (i.e., immunogenic epitopes) nor to the amino or carboxyl terminals. Antigenic epitope-bearing peptides and polypeptides of the invention are therefore useful to raise antibodies, including monoclonal antibodies, that bind specifically to a polypeptide of the invention. See, for instance, Wilson *et al.*, *Cell* 37:767-778 (1984) at 777.

[0121] Antigenic epitope-bearing peptides and polypeptides of the invention preferably contain a sequence of at least seven, more preferably at least nine and most preferably between about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention. Non-limiting examples of antigenic polypeptides or peptides that can be used to generate CK α -6-specific antibodies include: a polypeptide comprising amino acid residues from about Leu-26 to about Cys-34 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about Thr-36 to about Leu-45 in SEQ ID NO:2; a polypeptide comprising amino acid residues from about Thr-58 to about Leu-66 in SEQ ID NO:2; and a polypeptide comprising amino acid residues from about Pro-77 to about Val-84 in SEQ ID NO:2. These polypeptide fragments have been determined to bear antigenic epitopes of the CK α -6 protein by the analysis of the Jameson-Wolf antigenic index, as shown in Figure 3, above.

[0122] The epitope-bearing peptides and polypeptides of the invention may be produced by any conventional means. See, e.g., Houghten, R. A. (1985) "General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids." *Proc. Natl. Acad. Sci. USA* 82:5131-5135; this "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten *et al.* (1986).

[0123] Epitope-bearing peptides and polypeptides of the invention are used to induce antibodies according to methods well known in the art. See, for instance, Sutcliffe *et al.*, *supra*; Wilson *et al.*, *supra*; Chow, M. *et al.*, *Proc. Natl. Acad. Sci. USA* 82:910-914; and Bittle, F. J. *et al.*, *J. Gen. Virol.* 66:2547-2354 (1985). Immunogenic epitope-bearing peptides of the invention, i.e., those parts of a protein that elicit an antibody response when the whole protein is the immunogen, are identified according to methods known in the art. See, for instance, Geysen *et al.*, *supra*. Further still, U.S. Patent No. 5,194,392 to Geysen (1990) describes a general method of detecting or determining the sequence of monomers (amino acids or other

compounds) which is a topological equivalent of the epitope (i.e., a "mimotope") which is complementary to a particular paratope (antigen binding site) of an antibody of interest. More generally, U.S. Patent No. 4,433,092 to Geysen (1989) describes a method of detecting or determining a sequence of monomers which is a topographical equivalent of a ligand which is complementary to the ligand binding site of a particular receptor of interest. Similarly, U.S. Patent No. 5,480,971 to Houghten, R. A. et al. (1996) on Peralkylated Oligopeptide Mixtures discloses linear C1-C7-alkyl peralkylated oligopeptides and sets and libraries of such peptides, as well as methods for using such oligopeptide sets and libraries for determining the sequence of a peralkylated oligopeptide that preferentially binds to an acceptor molecule of interest. Thus, non-peptide analogs of the epitope-bearing peptides of the invention also can be made routinely by these methods.

Fusion Proteins

[0124] As one of skill in the art will appreciate, CK α -6 polypeptides of the present invention and the epitope-bearing fragments thereof described above can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. This has been shown, e.g., for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins (EP A 394,827; Traunecker et al., *Nature* 331:84-86 (1988)). Fusion proteins that have a disulfide-linked dimeric structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than the monomeric CK α -6 protein or protein fragment alone (Fountoulakis et al., *J. Biochem.* 270:3958-3964 (1995)).

Antibodies

[0125] CK α -6-protein specific antibodies for use in the present invention can be raised against the intact CK α -6 protein or an antigenic polypeptide fragment thereof, which may be presented together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse) or, if it is long enough (at least about 25 amino acids), without a carrier.

[0126] As used herein, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules as well as antibody fragments (such as, for example, Fab and F(ab')₂ fragments) which are capable of specifically binding to CK α -6 protein. Fab and F(ab')₂ fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation, and

may have less non-specific tissue binding of an intact antibody (Wahl *et al.*, *J. Nucl. Med.* 24:316-325 (1983)). Thus, these fragments are preferred.

[0127] The antibodies of the present invention may be prepared by any of a variety of methods. For example, cells expressing the CK α -6 protein or an antigenic fragment thereof can be administered to an animal in order to induce the production of sera containing polyclonal antibodies. In a preferred method, a preparation of CK α -6 protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity.

[0128] In the most preferred method, the antibodies of the present invention are monoclonal antibodies (or CK α -6 protein binding fragments thereof). Such monoclonal antibodies can be prepared using hybridoma technology (Köhler *et al.*, *Nature* 256:495 (1975); Köhler *et al.*, *Eur. J. Immunol.* 6:511 (1976); Köhler *et al.*, *Eur. J. Immunol.* 6:292 (1976); Hammerling *et al.*, in: *Monoclonal Antibodies and T-Cell Hybridomas*, Elsevier, N.Y., (1981) pp. 563-681). In general, such procedures involve immunizing an animal (preferably a mouse) with a CK α -6 protein antigen or, more preferably, with a CK α -6 protein-expressing cell. Suitable cells can be recognized by their capacity to bind anti-CK α -6 protein antibody. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56° C), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 μ g/ml of streptomycin. The splenocytes of such mice are extracted and fused with a suitable myeloma cell line. Any suitable myeloma cell line may be employed in accordance with the present invention; however, it is preferable to employ the parent myeloma cell line (SP2O), available from the American Type Culture Collection, Rockville, Maryland. After fusion, the resulting hybridoma cells are selectively maintained in HAT medium, and then cloned by limiting dilution as described by Wands *et al.* (*Gastroenterology* 80:225-232 (1981)). The hybridoma cells obtained through such a selection are then assayed to identify clones which secrete antibodies capable of binding the CK α -6 protein antigen.

[0129] Alternatively, additional antibodies capable of binding to the CK α -6 protein antigen may be produced in a two-step procedure through the use of anti-idiotypic antibodies. Such a method makes use of the fact that antibodies are themselves antigens, and that, therefore, it is possible to obtain an antibody which binds to a second antibody. In accordance

with this method, CK α -6-protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones which produce an antibody whose ability to bind to the CK α -6 protein-specific antibody can be blocked by the CK α -6 protein antigen. Such antibodies comprise anti-idiotypic antibodies to the CK α -6 protein-specific antibody and can be used to immunize an animal to induce formation of further CK α -6 protein-specific antibodies.

[0130] It will be appreciated that Fab and F(ab')₂ and other fragments of the antibodies of the present invention may be used according to the methods disclosed herein. Such fragments are typically produced by proteolytic cleavage, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')₂ fragments). Alternatively, CK α -6 protein-binding fragments can be produced through the application of recombinant DNA technology or through synthetic chemistry.

[0131] For *in vivo* use of anti-CK α -6 in humans, it may be preferable to use "humanized" chimeric monoclonal antibodies. Such antibodies can be produced using genetic constructs derived from hybridoma cells producing the monoclonal antibodies described above. Methods for producing chimeric antibodies are known in the art. See, for review, Morrison, *Science* 229:1202 (1985); Oi et al., *BioTechniques* 4:214 (1986); Cabilly et al., U.S. Patent No. 4,816,567; Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., *Nature* 312:643 (1984); Neuberger et al., *Nature* 314:268 (1985).

Central Nervous and Immune System-Related Disorders

Diagnosis

[0132] The present inventors have discovered that CK α -6 is expressed in central nervous system (CNS) and immune system tissues. For a number of CNS and immune system-related disorders, substantially altered (increased or decreased) levels of CK α -6 gene expression can be detected in CNS or immune system tissue or other cells or bodily fluids (e.g., sera, plasma, urine, synovial fluid or spinal fluid) taken from an individual having such a disorder, relative to a "standard" CK α -6 gene expression level, that is, the CK α -6 expression level in CNS or immune system tissues or bodily fluids from an individual not having the CNS or immune system disorder. Thus, the invention provides a diagnostic method useful during diagnosis of a

CNS or immune system disorder, which involves measuring the expression level of the gene encoding the CK α -6 protein in CNS or immune system tissue or other cells or body fluid from an individual and comparing the measured gene expression level with a standard CK α -6 gene expression level, whereby an increase or decrease in the gene expression level compared to the standard is indicative of a CNS or immune system disorder.

[0133] In particular, it is believed that certain tissues in mammals with cancer, particularly ovarian cancer, express significantly altered levels of the CK α -6 protein and mRNA encoding the CK α -6 protein when compared to a corresponding "standard" level. Further, it is believed that altered levels of the CK α -6 protein can be detected in certain body fluids (e.g., sera, plasma, urine, and spinal fluid) from mammals with such a cancer when compared to sera from mammals of the same species not having the cancer.

[0134] Thus, the invention provides a diagnostic method useful during diagnosis of a CNS or immune system disorder, including cancers of these systems and of the ovaries, which involves measuring the expression level of the gene encoding the CK α -6 protein in immune or skeletal system tissue or other cells or body fluid from an individual and comparing the measured gene expression level with a standard CK α -6 gene expression level, whereby an increase or decrease in the gene expression level compared to the standard is indicative of a disorder.

[0135] Where a diagnosis of a disorder in the CNS or immune system, including diagnosis of a tumor, has already been made according to conventional methods, the present invention is useful as a prognostic indicator, whereby patients exhibiting enhanced or depressed CK α -6 gene expression will experience a worse clinical outcome relative to patients expressing the gene at a level nearer the standard level.

[0136] By "assaying the expression level of the gene encoding the CK α -6 protein" is intended qualitatively or quantitatively measuring or estimating the level of the CK α -6 protein or the level of the mRNA encoding the CK α -6 protein in a first biological sample either directly (e.g., by determining or estimating absolute protein level or mRNA level) or relatively (e.g., by comparing to the CK α -6 protein level or mRNA level in a second biological sample). Preferably, the CK α -6 protein level or mRNA level in the first biological sample is measured or estimated and compared to a standard CK α -6 protein level or mRNA level, the standard being taken from a second biological sample obtained from an individual not having the disorder or being determined by averaging levels from a population of individuals not having a

disorder of the immune system. As will be appreciated in the art, once a standard CK α -6 protein level or mRNA level is known, it can be used repeatedly as a standard for comparison.

[0137] By "biological sample" is intended any biological sample obtained from an individual, body fluid, cell line, tissue culture, or other source which contains CK α -6 protein or mRNA. As indicated, biological samples include body fluids (such as sera, plasma, urine, synovial fluid and spinal fluid) which contain free CK α -6 protein, CNS, immune system tissue, and other tissue sources found to express complete mature CK α -6 or a CK α -6 receptor. Methods for obtaining tissue biopsies and body fluids from mammals are well known in the art. Where the biological sample is to include mRNA, a tissue biopsy is the preferred source.

[0138] The present invention is useful for diagnosis or treatment of various CNS and immune system-related disorders in mammals, preferably humans. Such disorders include tumors, cancers, particularly ovarian cancer and any dysregulation of immune cell function including, but not limited to, autoimmunity, arthritis, leukemias, lymphomas, immunosuppression, immunity, humoral immunity, inflammatory bowel disease, myelo suppression, and the like.

[0139] Total cellular RNA can be isolated from a biological sample using any suitable technique such as the single-step guanidinium-thiocyanate-phenol-chloroform method described in Chomczynski and Sacchi, *Anal. Biochem.* 162:156-159 (1987). Levels of mRNA encoding the CK α -6 protein are then assayed using any appropriate method. These include Northern blot analysis, S1 nuclease mapping, the polymerase chain reaction (PCR), reverse transcription in combination with the polymerase chain reaction (RT-PCR), and reverse transcription in combination with the ligase chain reaction (RT-LCR).

[0140] Assaying CK α -6 protein levels in a biological sample can occur using antibody-based techniques. For example, CK α -6 protein expression in tissues can be studied with classical immunohistological methods (Jalkanen, M., *et al.*, *J. Cell. Biol.* 101:976-985 (1985); Jalkanen, M., *et al.*, *J. Cell. Biol.* 105:3087-3096 (1987)). Other antibody-based methods useful for detecting CK α -6 protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (^{125}I , ^{121}I), carbon (^{14}C), sulfur (^{35}S), tritium (^3H), indium (^{112}In), and technetium ($^{99\text{m}}\text{Tc}$), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

[0141] In addition to assaying CK α -6 protein levels in a biological sample obtained from an individual, CK α -6 protein can also be detected *in vivo* by imaging. Antibody labels or markers for *in vivo* imaging of CK α -6 protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

[0142] A CK α -6 protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, ^{131}I , ^{112}In , $^{99\text{m}}\text{Tc}$), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously or intraperitoneally) into the mammal to be examined for immune system disorder. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of $^{99\text{m}}\text{Tc}$. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain CK α -6 protein. *In vivo* tumor imaging is described in S.W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments" (Chapter 13 in *Tumor Imaging: The Radiochemical Detection of Cancer*, S.W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982)).

Treatment

[0143] As noted above, CK α -6 polynucleotides and polypeptides are useful for diagnosis of conditions involving abnormally high or low expression of CK α -6 activities. Given the cells and tissues where CK α -6 is expressed as well as the activities modulated by CK α -6, it is readily apparent that a substantially altered (increased or decreased) level of expression of CK α -6 in an individual compared to the standard or "normal" level produces pathological conditions related to the bodily system(s) in which CK α -6 is expressed and/or is active.

[0144] It will also be appreciated by one of ordinary skill that, since the CK α -6 protein of the invention is a member of the chemokine family the mature secreted form of the protein may be released in soluble form from the cells which express the CK α -6 polypeptide. Therefore,

when CK α -6 is added from an exogenous source to cells, tissues or the body of an individual, the protein will exert its physiological activities on its target cells of that individual.

[0145] Therefore, it will be appreciated that conditions caused by a decrease in the standard or normal level of CK α -6 activity in an individual, particularly disorders of the CNS and immune system, can be treated by administration of CK α -6 polypeptide in the form of the mature protein. Thus, the invention also provides a method of treatment of an individual in need of an increased level of CK α -6 activity comprising administering to such an individual a pharmaceutical composition comprising an amount of an isolated CK α -6 polypeptide of the invention, effective to increase the CK α -6 activity level in such an individual.

[0146] Since CK α -6 lacks an ELR motif, it also should be an angiostatic factor rather than an angiogenic factor. In addition, since CK α -6 inhibits endothelial cell function, it will have a wide range of anti-inflammatory activities. CK α -6 may be employed as an anti-neovascularizing agent to treat solid tumors by stimulating the invasion and activation of host defense cells, e.g., cytotoxic T cells and macrophages and by inhibiting the angiogenesis of tumors. Those of skill in the art will recognize other non-cancer indications where blood vessel proliferation is not wanted. They may also be employed to enhance host defenses against resistant chronic and acute infections, for example, myobacterial infections via the attraction and activation of microbicidal leukocytes. CK α -6 may also be employed to inhibit T-cell proliferation by the inhibition of IL-2 biosynthesis for the treatment of T-cell mediated auto-immune diseases and lymphocytic leukemias. CK α -6 may also be employed to stimulate wound healing, both via the recruitment of debris clearing and connective tissue promoting inflammatory cells and also via its control of excessive TGF-mediated fibrosis. In this same manner, CK α -6 may also be employed to treat other fibrotic disorders, including liver cirrhosis, osteoarthritis and pulmonary fibrosis. CK α -6 also increases the presence of eosinophils which have the distinctive function of killing the larvae of parasites that invade tissues, as in schistosomiasis, trichinosis and ascariasis. It may also be employed to regulate hematopoiesis, by regulating the activation and differentiation of various hematopoietic progenitor cells, for example, to release mature leukocytes from the bone marrow following chemotherapy, i.e., in stem cell mobilization. CK α -6 may also be employed to treat sepsis. Also, by making a specific mutation to include an ELR motif, CK α -6-ELR will have angiogenic activities which are useful for treating all of the disease states where angiogenesis would be beneficial, i.e., to promote wound healing, re-vascularization of damaged limbs from injury or disease, and others

known to those of skill in the art. Polypeptides of the present invention may be useful in the prevention of HIV infection since it is known that chemokine receptors act as receptors for HIV infection.

Formulations

[0147] The CK α -6 polypeptide composition will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with CK α -6 polypeptide alone), the site of delivery of the CK α -6 polypeptide composition, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" of CK α -6 polypeptide for purposes herein is thus determined by such considerations.

[0148] As a general proposition, the total pharmaceutically effective amount of CK α -6 polypeptide administered parenterally per dose will be in the range of about 1 μ g/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the CK α -6 polypeptide is typically administered at a dose rate of about 1 μ g/kg/hour to about 50 μ g/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

[0149] Pharmaceutical compositions containing the CK α -6 of the invention may be administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, drops or transdermal patch), buccally, or as an oral or nasal spray. By "pharmaceutically acceptable carrier" is meant a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

[0150] The CK α -6 polypeptide is also suitably administered by sustained-release systems. Suitable examples of sustained-release compositions include semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules. Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of L-glutamic acid and gamma-

ethyl-L-glutamate (Sidman, U. et al., *Biopolymers* 22:547-556 (1983)), poly (2- hydroxyethyl methacrylate) (R. Langer et al., *J. Biomed. Mater. Res.* 15:167-277 (1981), and R. Langer, *Chem. Tech.* 12:98-105 (1982)), ethylene vinyl acetate (R. Langer et al., Id.) or poly-D- (-)-3-hydroxybutyric acid (EP 133,988). Sustained-release CK α -6 polypeptide compositions also include liposomally entrapped CK α -6 polypeptide. Liposomes containing CK α -6 polypeptide are prepared by methods known per se: DE 3,218,121; Epstein et al., *Proc. Natl. Acad. Sci. (USA)* 82:3688-3692 (1985); Hwang et al., *Proc. Natl. Acad. Sci. (USA)* 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal CK α -6 polypeptide therapy.

[0151] For parenteral administration, in one embodiment, the CK α -6 polypeptide is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to polypeptides.

[0152] Generally, the formulations are prepared by contacting the CK α -6 polypeptide uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

[0153] The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as

polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, manose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

[0154] The CK α -6 polypeptide is typically formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of CK α -6 polypeptide salts.

[0155] CK α -6 polypeptide to be used for therapeutic administration must be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutic CK α -6 polypeptide compositions generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

[0156] CK α -6 polypeptide ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous CK α -6 polypeptide solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized CK α -6 polypeptide using bacteriostatic Water-for-Injection.

[0157] The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the polypeptides of the present invention may be employed in conjunction with other therapeutic compounds.

Agonists and Antagonists - Assays and Molecules

[0158] The invention also provides a method of screening compounds to identify those which enhance or block the action of CK α -6 on cells, such as its interaction with CK α -6-binding molecules such as receptor molecules. An agonist is a compound which increases the

natural biological functions of CK α -6 or which functions in a manner similar to CK α -6, while antagonists decrease or eliminate such functions.

[0159] In another aspect of this embodiment the invention provides a method for identifying a receptor protein or other ligand-binding protein which binds specifically to a CK α -6 polypeptide. For example, a cellular compartment, such as a membrane or a preparation thereof, may be prepared from a cell that expresses a molecule that binds CK α -6. The preparation is incubated with labeled CK α -6 and complexes of CK α -6 bound to the receptor or other binding protein are isolated and characterized according to routine methods known in the art. Alternatively, the CK α -6 polypeptide may be bound to a solid support so that binding molecules solubilized from cells are bound to the column and then eluted and characterized according to routine methods.

[0160] In the assay of the invention for agonists or antagonists, a cellular compartment, such as a membrane or a preparation thereof, may be prepared from a cell that expresses a molecule that binds CK α -6, such as a molecule of a signaling or regulatory pathway modulated by CK α -6. The preparation is incubated with labeled CK α -6 in the absence or the presence of a candidate molecule which may be a CK α -6 agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand. Molecules which bind gratuitously, i.e., without inducing the effects of CK α -6 on binding the CK α -6 binding molecule, are most likely to be good antagonists. Molecules that bind well and elicit effects that are the same as or closely related to CK α -6 are agonists.

[0161] CK α -6-like effects of potential agonists and antagonists may be measured, for instance, by determining activity of a second messenger system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of CK α -6 or molecules that elicit the same effects as CK α -6. Second messenger systems that may be useful in this regard include but are not limited to AMP guanylate cyclase, ion channel or phosphoinositide hydrolysis second messenger systems.

[0162] Another example of an assay for CK α -6 antagonists is a competitive assay that combines CK α -6 and a potential antagonist with membrane-bound CK α -6 receptor molecules or recombinant CK α -6 receptor molecules under appropriate conditions for a competitive inhibition assay. CK α -6 can be labeled, such as by radioactivity, such that the number of CK α -6 molecules bound to a receptor molecule can be determined accurately to assess the effectiveness of the potential antagonist.

[0163] Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as a receptor molecule, without inducing CK α -6-induced activities, thereby preventing the action of CK α -6 by excluding CK α -6 from binding.

[0164] Other potential antagonists include antisense molecules. Antisense technology can be used to control gene expression through antisense DNA or RNA or through triple-helix formation. Antisense techniques are discussed, for example, in Okano, *J. Neurochem.* 56: 560 (1991); "Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression." CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance Lee *et al.*, *Nucleic Acids Research* 6: 3073 (1979); Cooney *et al.*, *Science* 241: 456 (1988); and Dervan *et al.*, *Science* 251: 1360 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of CK α -6. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into CK α -6 polypeptide. The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of CK α -6 protein.

[0165] The agonists and antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as described above.

[0166] The antagonists may be employed for instance to inhibit the chemotaxis and activation of macrophages and their precursors, and of neutrophils, basophils, B lymphocytes and some T-cell subsets, e.g., activated and CD8 cytotoxic T cells and natural killer cells, in certain auto-immune and chronic inflammatory and infective diseases. Examples of auto-immune diseases include multiple sclerosis, and insulin-dependent diabetes. The antagonists may also be employed to treat infectious diseases including silicosis, sarcoidosis, idiopathic pulmonary fibrosis by preventing the recruitment and activation of mononuclear phagocytes. They may also be employed to treat idiopathic hyper-eosinophilic syndrome by preventing eosinophil production and migration. Endotoxic shock may also be treated by the antagonists

by preventing the migration of macrophages and their production of the human chemokine polypeptides of the present invention. The antagonists may also be employed for treating atherosclerosis, by preventing monocyte infiltration in the artery wall. The antagonists may also be employed to treat histamine-mediated allergic reactions and immunological disorders including late phase allergic reactions, chronic urticaria, and atopic dermatitis by inhibiting chemokine-induced mast cell and basophil degranulation and release of histamine. IgE-mediated allergic reactions such as allergic asthma, rhinitis, and eczema may also be treated. The antagonists may also be employed to treat chronic and acute inflammation by preventing the attraction of monocytes to a wound area. They may also be employed to regulate normal pulmonary macrophage populations, since chronic and acute inflammatory pulmonary diseases are associated with sequestration of mononuclear phagocytes in the lung. Antagonists may also be employed to treat rheumatoid arthritis by preventing the attraction of monocytes into synovial fluid in the joints of patients. Monocyte influx and activation plays a significant role in the pathogenesis of both degenerative and inflammatory arthropathies. The antagonists may be employed to interfere with the deleterious cascades attributed primarily to IL-1 and TNF, which prevents the biosynthesis of other inflammatory cytokines. In this way, the antagonists may be employed to prevent inflammation. The antagonists may also be employed to inhibit prostaglandin-independent fever induced by chemokines. The antagonists may also be employed to treat cases of bone marrow failure, for example, aplastic anemia and myelodysplastic syndrome. The antagonists may also be employed to treat asthma and allergy by preventing eosinophil accumulation in the lung. The antagonists may also be employed to treat subepithelial basement membrane fibrosis which is a prominent feature of the asthmatic lung. Antibodies against CK α -6 may be employed to bind to and inhibit CK α -6 activity to treat ARDS, by preventing infiltration of neutrophils into the lung after injury. Any of the above antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as hereinafter described.

Gene Mapping

[0167] The nucleic acid molecules of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal

Examples

Example 1: Cloning and Expression of CK α -6 in *E. coli*

[0174] The present invention includes an expression vector comprising phage operator and promoter elements operatively linked to a polynucleotide of the present invention, called pHE4a. (ATCC Accession Number 209645, deposited on February 25, 1998.) This vector contains: 1) a neomycinphosphotransferase gene as a selection marker, 2) an *E. coli* origin of replication, 3) a T5 phage promoter sequence, 4) two lac operator sequences, 5) a Shine-Delgarno sequence, and 6) the lactose operon repressor gene (lacIq). The origin of replication (oriC) is derived from pUC19 (LTI, Gaithersburg, MD). The promoter sequence and operator sequences were made synthetically.

[0175] DNA can be inserted into the pHEa by restricting the vector with NdeI and XbaI, BamHI, XhoI, or Asp718, running the restricted product on a gel, and isolating the larger fragment (the stuffer fragment should be about 310 base pairs). The DNA insert is generated according to the PCR protocol described in Example 1, using PCR primers having restriction sites for NdeI (5' primer) and XbaI, BamHI, XhoI, or Asp718 (3' primer). The PCR insert is gel purified and restricted with compatible enzymes. The insert and vector are ligated according to standard protocols. Preferred primers for subcloning CK α -6 have the following amino acid sequences:

[0176] 5' primer: 5'-GCGCCATATGCGGGTGGTTTGGGG-3' (SEQ ID NO:15) (the NdeI site is underlined);

[0177] 3' primer: 5'-CGCGAATTCTTAAACATTCTTACCAGG-3' (SEQ ID NO:16) (the EcoRI site is underlined).

[0178] The ligation mixture is used to transform the *E. coli* strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kanr). Transformants are identified by their ability to grow on LB plates and neomycin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

[0179] Clones containing the desired constructs are grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 μ g/ml) and kanamycin (25

[illegible]

[0181] The following alternative method can be used to purify a polypeptide expressed in E coli when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10°C.

[0183] The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000 xg for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

50

discarded and the polypeptide containing supernatant is incubated at 4°C overnight to allow further GuHCl extraction.

[0185] Following high speed centrifugation (30,000 xg) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4°C without mixing for 12 hours prior to further purification steps.

[0186] To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 µm membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perseptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

[0187] Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perseptive Biosystems) and weak anion (Poros CM-20, Perseptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A280 monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

[0188] The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Commassie blue stained 16% SDS-PAGE gel when 5 µg of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 2: Cloning and Expression of CK α -6 protein in a Baculovirus Expression System

[0189] In this illustrative example, the plasmid shuttle vector pA2 is used to insert the cloned DNA encoding complete protein, including its naturally associated secretory signal (leader) sequence, into a baculovirus to express the mature CK α -6 protein, using standard methods as described in Summers *et al.*, *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures*, Texas Agricultural Experimental Station Bulletin No. 1555 (1987). This expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites such as BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak *Drosophila* promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

[0190] Many other baculovirus vectors could be used in place of the vector above, such as pAc373, pVL941 and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow *et al.*, *Virology* 170:31-39 (1989).

[0191] The cDNA sequence encoding the full length CK α -6 protein in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence shown in SEQ ID NO:2, is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene. The 5' primer has the sequence 5' CGCGGATCCGCCATCATGCA GAGGCCCTTCCTC 3' (SEQ ID NO:17) containing the underlined BamHI restriction enzyme site, an efficient signal for initiation of translation in eukaryotic cells, as described by Kozak, M., *J. Mol. Biol.* 196:947-950 (1987). The 3' primer has the sequence 5' GCGTCTAGATCAAACATTCTTACCAGG 3' (SEQ ID NO:18) containing the underlined XbaI restriction site. Those of ordinary skill in the art will recognize that other primers could be used to generate nucleic acid fragments encoding shorter polypeptides.

[0192] The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with BamHI and XbaI and again is purified on a 1% agarose gel.

[0193] The plasmid is digested with the restriction enzymes BamHI and XbaI and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.).

[0194] Fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) cells are transformed with the ligation mixture and spread on culture plates. Bacteria are identified that contain the plasmid with the human CK α -6 gene by digesting DNA from individual colonies using BamHI and XbaI and then analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing. This plasmid is designated herein pA2GPCK α -6.

[0195] Five μ g of the plasmid pA2GPCK α -6 is co-transfected with 1.0 μ g of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner et al., *Proc. Natl. Acad. Sci. USA* 84: 7413-7417 (1987). One μ g of BaculoGold™ virus DNA and 5 μ g of the plasmid pA2GPCK α -6 are mixed in a sterile well of a microtiter plate containing 50 μ l of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards, 10 μ l Lipofectin plus 90 μ l Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27° C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27° C for four days.

[0196] After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, *supra*. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by

Life Technologies Inc., Gaithersburg, page 9-10). After appropriate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 μ l of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4° C. The recombinant virus is called V-CK α -6.

[0197] To verify the expression of the CK α -6 gene Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-CK α -6 at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, MD). After 42 hours, 5 μ Ci of ³⁵S-methionine and 5 μ Ci ³⁵S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

[0198] Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the mature form or extracellular domain of the CK α -6 protein and thus the cleavage point and length of the naturally associated secretory signal peptide.

Example 3: Cloning and Expression of CK α -6 in Mammalian Cells

[0199] A typical mammalian expression vector contains the promoter element, which mediates the initiation of transcription of mRNA, the protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription can be achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLV1, HIV1 and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter). Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and

pBC12MI (ATCC 67109). Mammalian host cells that could be used include, human HeLa, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

[0200] Alternatively, the gene can be expressed in stable cell lines that contain the gene integrated into a chromosome. The co-transfection with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transfected cells.

[0201] The transfected gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful to develop cell lines that carry several hundred or even several thousand copies of the gene of interest. Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., *Biochem J.* 227:277-279 (1991); Bebbington et al., *Bio/Technology* 10:169-175 (1992)). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

[0202] The expression vectors pC1 and pC4 contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen et al., *Molecular and Cellular Biology*, 438-447 (March, 1985)) plus a fragment of the CMV-enhancer (Boshart et al., *Cell* 41:521-530 (1985)). Multiple cloning sites, e.g., with the restriction enzyme cleavage sites BamHI, XbaI and Asp718, facilitate the cloning of the gene of interest. The vectors contain in addition the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene.

Example 3(a): Cloning and Expression in COS Cells

[0203] The expression plasmid, pCK α -6HA, is made by cloning a portion of the cDNA encoding the extracellular domain of the CK α -6 protein into the expression vector pcDNAI/Amp or pcDNAIII (which can be obtained from Invitrogen, Inc.).

[0204] The expression vector pcDNAI/amp contains: (1) an *E. coli* origin of replication effective for propagation in *E. coli* and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron; (5) several codons encoding a hemagglutinin fragment (i.e., an "HA" tag to facilitate purification) followed by a termination codon and polyadenylation signal arranged so that a cDNA can be conveniently placed under expression control of the CMV promoter and operably linked to the

SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson et al., *Cell* 37: 767 (1984). The fusion of the HA tag to the target protein allows easy detection and recovery of the recombinant protein with an antibody that recognizes the HA epitope. pcDNAIII contains, in addition, the selectable neomycin marker.

[0205] A DNA fragment encoding the soluble extracellular domain of the CK α -6 polypeptide is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter. The plasmid construction strategy is as follows. The CK α -6 cDNA of the deposited clone is amplified using primers that contain convenient restriction sites, much as described above for construction of vectors for expression of CK α -6 in *E. coli*. Suitable primers include the following, which are used in this example. The 5' primer, containing the underlined BamHI site, a Kozak sequence, and an AUG start codon, has the following sequence: 5' CGCGGATCCGCCATCATGCAGAGGCCCTTCCTC 3' (SEQ ID NO:17). The 3' primer, containing the underlined XbaI site has the following sequence: 5' GCGTCTAGATCAAACATTCTTACCAGG 3' (SEQ ID NO:18).

[0206] The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with BamHI and XbaI and then ligated. The ligation mixture is transformed into *E. coli* strain SURE (available from Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037), and the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis or other means for the presence of the fragment encoding the extracellular domain of the CK α -6 polypeptide

[0207] For expression of recombinant CK α -6, COS cells are transfected with an expression vector, as described above, using DEAE-DEXTRAN, as described, for instance, in Sambrook et al., *Molecular Cloning: a Laboratory Manual*, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989). Cells are incubated under conditions for expression of CK α -6 by the vector.

[0208] Expression of the CK α -6-HA fusion protein is detected by radiolabeling and immunoprecipitation, using methods described in, for example Harlow et al., *Antibodies: A Laboratory Manual*, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988). To this end, two days after transfection, the cells are labeled by incubation in

a) b) c) d) e) f) g) h) i) j) k) l) m) n) o) p) q) r) s) t) u) v) w) x) y) z)

[0209] The vector pC4 is used for the expression of CK α -6 polypeptide in this example. Plasmid pC4 is a derivative of the plasmid pSV2-dhfr (ATCC Accession No. 37146). The plasmid contains the mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary- or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selected by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with the chemotherapeutic agent methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented (see, e.g., Alt, F. W., Kellems, R. M., Bertino, J. R., and Schimke, R. T., 1978, *J. Biol. Chem.* 253:1357-1370, Hamlin, J. L. and Ma, C. 1990, *Biochem. et Biophys. Acta*, 1097:107-143, Page, M. J. and Sydenham, M. A. 1991, *Biotechnology* 9:64-68). Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene, it is usually co-amplified and over-expressed. It is known in the art that this approach may be used to develop cell lines carrying more than 1,000 copies of the amplified gene(s). Subsequently, when the methotrexate is withdrawn, cell lines are obtained which contain the amplified gene integrated into one or more chromosome(s) of the host cell.

high efficiency promoters can also be used for the expression, e.g., the human β -actin promoter, the SV40 early or late promoters or the long terminal repeats from other retroviruses, e.g., HIV and HTLV. Clontech's Tet-Off and Tet-On gene expression systems and similar systems can be used to express the CK α -6 polypeptide in a regulated way in mammalian cells (Gossen, M., & Bujard, H. 1992, *Proc. Natl. Acad. Sci. USA* 89:5547-5551). For the polyadenylation of the mRNA other signals, e.g., from the human growth hormone or globin genes can be used as well. Stable cell lines carrying a gene of interest integrated into the chromosomes can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

[0211] The plasmid pC4 is digested with the restriction enzymes BamHI and XbaI and then dephosphorylated using calf intestinal phosphatase by procedures known in the art. The vector is then isolated from a 1% agarose gel.

[0212] The DNA sequence encoding the extracellular domain of the CK α -6 polypeptide is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the desired portion of the gene. The 5' and 3' primers are the same as those used in Example 3(a) above.

[0213] The amplified fragment is digested with the endonucleases BamHI and XbaI and then purified again on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

[0214] Chinese hamster ovary cells lacking an active DHFR gene are used for transfection. Five μ g of the expression plasmid pC4 is cotransfected with 0.5 μ g of the plasmid pSVneo using lipofectin (Felgner et al., *supra*). The plasmid pSV2-neo contains a dominant selectable marker, the *neo* gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of

methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 μ M, 2 μ M, 5 μ M, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 μ M. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 4: Tissue distribution of CK α -6 mRNA expression

[0215] Northern blot analysis is carried out to examine CK α -6 gene expression in human tissues, using methods described by, among others, Sambrook *et al.*, cited above. A cDNA probe containing the entire nucleotide sequence of the CK α -6 protein (SEQ ID NO:1) is labeled with 32 P using the *rediprime*TM DNA labeling system (Amersham Life Science), according to manufacturer's instructions. After labeling, the probe is purified using a CHROMA SPIN-100TM column (Clontech Laboratories, Inc.), according to manufacturer's protocol number PT1200-1. The purified labeled probe is then used to examine various human tissues for CK α -6 mRNA.

[0216] Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) are obtained from Clontech and are examined with the labeled probe using ExpressHybTM hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1, incorporated herein by reference. Following hybridization and washing, the blots are mounted and exposed to film at -70° C overnight, and films developed according to standard procedures.

[0217] It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

[0218] The entire disclosure of all publications (including patents, patent applications, journal articles, laboratory manuals, books, or other documents) cited herein are hereby incorporated by reference. Further, the sequence listing, both the hard copy and CRF, submitted herewith is hereby incorporated herein by reference.